The CORE datasets from Large and Yeager

Stephen.Griffies@noaa.gov, Michael.Winton@noaa.gov, and Bonnie.Samuels@noaa.gov
NOAA/Geophysical Fluid Dynamics Laboratory
Princeton, USA
Manuscript last edited May 23, 2008

Abstract

These notes describe the release of the *CORE forcing* developed by Bill Large and Steve Yeager at NCAR for global ocean-ice modelling. CORE.v1 includes both an interannually varying forcing (IAF) from 1958-2004, and a normal year forcing (NYF) derived from the interannual forcing. CORE.v1 is detailed in the technical report Large and Yeager (2004). CORE.v2 from Large and Yeager was released during May 2008 and it is documented in Large and Yeager (2008). This second release updates just the interannually varying forcing, now extending from 1958-2006. Both CORE.v1 and CORE.v2 are supported at NCAR and GFDL for use in studying global ocean-ice dynamics.

These release notes, the CORE datasets, support code, and further documentation can be downloaded from

`http://data1.gfdl.noaa.gov/nomads/forms/mom4/CORE.html`

1 Introduction

This document describes the release of the CORE datatsets. It was written by GFDL scientists in support of the CLIVAR Working Group for Ocean Model Development (WGOMD) Coordinated Ocean Reference Experiments (CORE). A key element facilitating the existence of CORE is the work of Bill Large and Steve Yeager. They developed algorithms that modify or "correct" various atmospheric datasets, so that these datasets can be used to derive surface fluxes needed to integrate global ocean-ice models. The Large and Yeager (2004) algorithms have been used to produce Version 1 of the interannually varying forcing (IAF) CORE.v1 for years 1958-2004, as well as a normal year forcing (NYF) derived from the interannual forcing. We refer to these "corrected" datasets as CORE-IAF.v1 and CORE-NYF.v1. Large and Yeager (2008) present updated algorithms for the interannual dataset 1958-2006, thus producing CORE-IAF.v2. CORE-NYF.v2 has not yet been derived.

At NCAR, the Large and Yeager corrections are applied to the uncorrected datasets during the runtime of a particular ocean-ice simulation. At GFDL, corrections are applied to the uncorrected datasets to produce a corrected dataset, which is then used to integrate the ocean-ice models. The NCAR strategy is preferred when developing the correction algorithms. Once a final suite of corrections has been derived, it is sensible to work with the corrected datasets. This web site provides options for both approaches.

2 The CORE dataset web page

We summarize here the contents of the CORE dataset web page

`http://data1.gfdl.noaa.gov/nomads/forms/mom4/CORE.html`

2.1 Datasets

The CORE dataset web page contains the following datasets.

- Version 1 datasets

    - Uncorrected Normal Year Forcing (unCNYF.v1)
    - Uncorrected Interannual Forcing (unCIAF.v1)
    - Corrected Normal Year Forcing (CORE-NYF.v1)
    - Corrected Interannual Forcing (CORE-IAF.v1)

- Version 2 datasets

  - Uncorrected Interannual Forcing (unCIAF.v2)
  - Corrected Interannual Forcing (CORE-IAF.v2)

Each of the above datasets contain the following fields on a spherical grid of 192 longitude points and 94 latitude points (T62 atmospheric grid):

- annual mean river runoff (1 time step for full dataset)

- monthly varying precipitation (12 time steps per year)

- daily varying shortwave and longwave (365 time steps per year–no diurnal cycle and no leap years),

- six-hourly varying 10m temperature, density, humidity, zonal velocity, meridional velocity, and sea level pressure (4×365 time steps per year–no leap years).

The unCIAF.v1 and unCIAF.v2 datasets are identical for years 1958-2003. However, both the NCEP fields and the GISS radiation data for 2004 were modified subsequent to the release of unCIAF.v1. Hence, year 2004 in unCIAF.v2 includes the changes. The precipitation for year 2004 has also been updated for months 10-12 in unCIAF.v2.

2.2 Support code and documentation

Besides the present set of notes and the datasets, the CORE web page also contains the following files.

- Version 1 support files

  - Large and Yeager (2004): This report details both the uncorrected and corrected data sets used to produce the forcing fields. In particular, it provides an atlas of the fluxes produced when using Reynolds SSTs and the NCAR bulk formula to compute fluxes from the atmospheric state.
  - Griffies et al. (2008): (CORE_NYFv1.pdf): This manuscript documents seven global ocean-ice models run with CORE-NYF.v1 for 500 years.
  - The Fortran code `advance.f90` provided by NCAR corrects the raw data. This code may be of use for those who compute the data corrections as the model integrates.
  - The Ferret code `make_data.csh` provided by GFDL implements the algorithms from `advance.f90` in a Ferret script.
  - The Fortran code `ncar_ocean_fluxes.f90` provided by GFDL computes the NCAR exchange coefficients recommended for use in CORE. This is the same as the Version 1 release.
  - The sea surface salinity restoring file `PHC2_salx.nc` provided by NCAR for use in computing a restoring salt or fresh water flux with CORE.

- Version 2 support files

  - The Large and Yeager (2008) preprint which documents the CORE-IAF.v2.
  - `README_COREv2` is a README file for the release of CORE-v2.0 from NCAR.
  - The Fortran code `datm_physTN460.F90` provided by NCAR corrects the raw data in the case that a user wishes to make the corrections during a run (online) rather than prior to the run.
  - The NetCDF file `tn460nyf.correction_factors.T62.121007.nc` provides the correction factors that are applied to the uncorrected datasets if users wish to run with `datm_physTN460.F90`.
  - The Ferret code `make_data_CIAFv2.2008_04_22.csh` provided by GFDL implements the algorithms from `datm_physTN460.F90` in a Ferret script.

- The Fortran code `ncar_ocean_fluxes.f90` provided by GFDL computes the NCAR exchange coefficients recommended for use in CORE. This is the same as the Version 1 release.

- The sea surface salinity restoring file `PHC2_salx.nc` provided by NCAR for use in computing a restoring salt or fresh water flux with CORE. This is the same as the Version 1 release.

We provide both the uncorrected and corrected forcing fields for two reasons.

- The user may wish to run simulations as at NCAR whereby corrections are applied to the uncorrected fields at runtime by using `advance.f90` for CORE.v1 or `datm_physTN460.F90` for CORE.v2. This procedure facilitates further refinement to the corrections without needing to generate a new "corrected" dataset.

- At GFDL, we perform corrections prior to runtime using the above Ferret script.

3 Reasons favouring the use of CORE forcing

The release of the CORE.v1 and CORE.v2 datasets by Large and Yeager provides the global ocean climate modeling community with an important advance in our ability to integrate ocean-ice models without a fully coupled atmospheric GCM. This advance builds in many ways on an earlier effort by Röske (2001) for a Pilot-Ocean Model Intercomparison Project (POMIP). There are generally various datasets that can be used for running coupled ocean and sea ice models. However, we prefer the CORE.v1 and CORE.v2 data for the following reasons.

- The data, which combines reanalysis with satellite data, has advantages over that based solely on reanalysis (Röske, 2001, was based solely on ECMWF). Advantages are discussed in Large and Yeager (2004).

- Both normal year *and* interannual data are provided in CORE.v1. Many researchers find the use of interannually varying data to be more interesting, since it better facilitates comparisons of model simulations with ocean observations.

- The datasets are documented and supported by NCAR, with extensive refinement as more data is gathered. GFDL has agreed to support the release of corrected versions and to document this web page to assist those who wish to use the datasets. Future releases of this data can be expected as improvements are made to the data products and to our understanding of their biases.

- The CLIVAR WGOMD has recomended the Large and Yeager derived datasets for use in various model comparison efforts, such as that documented in Griffies et al. (2008).

4 Comments on the data and experimental methods

We now present some comments on particular aspects of using the CORE datasets. Any recommendations provided here are subject to modifications as different modeling groups gain experience with this forcing. Note that Griffies et al. (2008) documents seven global ocean-ice models employing the CORE-NYF.v1 forcing for 500 years. Experience with the interannual varying data at GFDL is minimal, as this data has only recently been developed. However, NCAR has been working with earlier versions of this dataset for some years, with the papers by Doney et al. (2003) and Doney et al. (2007) providing guidance for its use. Furthermore, with the release of CORE-IAF.v2, we anticipate much more use of this dataset at GFDL and elsewhere.

4.1 Initial conditions and experiment duration

We have generally run the CORE-NYF.v1 simulations for at least 100 years, with 500 years preferred for reasons discussed in Griffies et al. (2008). The models are initialized in the ocean with Levitus, and sea ice models are generally begun with a state taken from an earlier simulation.

Initialization and experiment length for the IAF simulations remains a research question. Some groups have proposed starting from Levitus. But the question remains how long to run for a spin-up phase. Some groups have proposed to run at least two realizations of the interannually varying forcing for spin-up, with an analysis period in the third realization. Doney et al. (2003) document some experience addressing this question.

4.2 Interannual forcing without leap-years

The interannual forcing fields in CORE-IAF.v1 and CORE-IAF.v2 do not contain leap-years. That is, each year has the same length of 365 days. This limitation may introduce some difficulties for those using the data for reanalysis efforts. However, the decision was made by NCAR to jettison the leap-years since many researchers find this to be more convenient given their software infrastructure.

4.3 Padding of years for the IAF

The IAF datasets provided on these web pages is split into individual years. There is no overlap. The transition from one year to another is a detail that is left to the respective modellers, as it is a function of the modeller's time interpolation code. At GFDL, we pad the corrected IAF data with a day on each side of the year boundary in order to smoothly time interpolate from one year to another.

4.4 Air density and sea level pressure

The subroutine `ncar_ocean_fluxes.f90` computes the exchange coefficients for momentum, evaporation, and sensible heat according to the equations documented in Large and Yeager (2004) (see their Section 2.1). After computing the exchange coefficients, the model computes air-sea fluxes based on equations (4a)-(4d) in Large and Yeager (2004). This calculation requires the air density. There are three ways to get this density, each of which result in rather small differences.

- The air density at 10m is provided in the uncorrected fields for version 2 of the IAF. Large and Yeager present no corrections to this field, so it can be used in CORE-IAF.v2.

- One may set air density to a constant $1.22 \mathrm{kg\,m^{-3}}$ (see Section 4.1 of Large and Yeager (2004)).

- One may use the sea level pressure provided in the CORE datasets, and then use the ideal gas law to compute the air density.

The preferred method depends on the structure of the flux computation code that each modeler maintains. At GFDL, we use the sea level pressure and ideal gas law, so we do not make use of the 10m air density dataset.

4.5 Radiative heating

Radiative heating is provided from the shortwave and longwave datasets. The shortwave and longwave datasets represent *downwelling* radiation. The *net* shortwave radiation $Q_{\mathrm{SW\ net}}$ transferred into the ocean is a function of the albedo as shown by equation (11) in Large and Yeager (2004). As discussed in Section 3.2 of Large and Yeager (2008), a latitudinally dependent albedo is used to compute the net shortwave in CORE-IAF.v2.

The net longwave radiation transferred into the ocean is given by the downwelling longwave radiation minus the loss of heat associated with re-radiation to the atmosphere as given by the Stefan-Boltzmann formulae $\sigma\,T^4$ as shown by equation (12) in Large and Yeager (2004).

The CORE datasets provide a single shortwave radiation field. However, many ocean optics models make use of four different partitions of this shortwave field: visible direct, visible diffuse, infrared direct, and infrared diffuse. NCAR recommends the following frabricated downward shortwave components for the purpose of mimicing a more complete atmospheric radiation model:

$$Q_{\text{visible direct}} = 0.28 \, Q_{\text{SW net}} \tag{1}$$
$$Q_{\text{IR direct}} = 0.31 \, Q_{\text{SW net}} \tag{2}$$
$$Q_{\text{visible diffuse}} = 0.24 \, Q_{\text{SW net}} \tag{3}$$
$$Q_{\text{IR diffuse}} = 0.17 \, Q_{\text{SW net}} \tag{4}$$

4.6 Surface temperature forcing

There is generally no restoring to surface temperature. Instead, turbulent heat fluxes are derived from the NCAR bulk formulae using the model SST and the 10m atmospheric fields, and radiative heating is provided by shortwave and longwave fluxes.

We initially tried to use the GFDL bulk formulae in our CORE-NYF.v1 simulations. However, the fluxes produced from the two bulk formulae are quite distinct when running with observed SSTs. In particular, the wind stresses are larger with the GFDL formulation (which follows ECMWF) and the latent heat fluxes are larger with the NCAR formulation. The differences have been traced to differences in the neutral transfer coefficients (roughness lengths). As the forcing datasets developed using the NCAR bulk formulae, we recommend using the same bulk formulae for CORE experiments.

We originally went into the NCAR/GFDL comparison thinking that the bulk formulae differences should lead to minor differences in the fluxes. However, the GFDL formulae is somewhat different than NCAR's. The resulting flux differences were too large to ignore, with the goal being to run the models with the same forcing when the SSTs were the same.

4.7 Properly referenced meteorological data

Models should use properly referenced meteorological data consistent with what the bulk formulae expect. Reanalysis meteorological data is commonly distributed at 2m while oceanic turbulent transfer schemes often require 10m data. For accuracy, it is essential that the data be re-referenced to 10m. The re-referencing algorithm and the flux calculation algorithm are closely related. So, one should re-reference using a scheme that is compatible with the flux scheme.

4.8 Same treatment of saltwater vapor pressure

Models should use the same treatment of saltwater vapor pressure. The vapor pressure over seawater is about 2% less than that over fresh water. This difference is not negligible compared to the 20% subsaturation of marine air that drives evaporation. Consequently, the effect should be included in all models participating in a comparison.

4.9 High frequency meteorological data

It is desirable to use high frequency meteorological data. A one month run of an AMIP model was used to explore the flux errors associated with averaged meteorological inputs. With daily winds, temperatures, and humidities, latent heat fluxes are under estimated broadly over the winter storm track band by some 10's of W/m2. There was also a smaller underestimate located in the summer storm track band. Experiments that refined the temporal resolution of the flux inputs individually showed that high frequency winds are most important for reducing the error but temperature and humidity frequency also contribute. When all inputs are given at 6 hourly frequency, the global RMS error is about 1 W/m2 versus near 8 W/m2 for daily inputs.

4.10 River runoff

The river runoff data has only a single time step as it represents annual mean runoff. This data has been spread out from the river mouths in a manner used by NCAR for their climate models. This approach is thought to account for some unresolved mixing that occurs at river mouths in Nature. We provide a remapping scheme which will take the river data and map onto a new grid, so long as the new grid is logically rectangular (such as the GFDL tripolar grid). GFDL can provide some assistance with this remapping if you have problems. Note that if modelers choose their own specification for runoff, perhaps with a seasonal cycle, we recommend that a correction be made to keep the total annual flux of runoff similar to the value in the Large and Yeager (2004) dataset in order to facilitate comparisons.

4.11 Salinity restoring

An issue for comparisons is the strength of the salinity restoring. Relatively strong salinity restoring, analogous to the effective restoring of SSTs, will reduce drift. However, salinity restoring has no physical basis, and so it is desirable to use the weakest possible restoring. A weak restoring also has the benefit of allowing increased variability in the surface salinity and deep circulation.

Unfortunately, when the salinity restoring and effective temperature restoring timescales are very different, the experiment becomes analogous to a mixed boundary condition experiment. The ability of mixed boundary conditions to represent the adjustment of the ocean in the coupled system has been called into question. In particular, mixed boundary condition experiments with strong temperature restoring have been shown to be excessively susceptible to the polar halocline catastrophe, in which a fresh cap develops in high latitudes and shuts down overturning (Zhang et al., 1993).

The effective temperature restoring determined by numerically linearizing the CORE thermal boundary condition is quite strong, yielding piston velocities around 1-2 m/day. The salinity restoring strength chosen for a comparison between NCAR and GFDL simulations with the normal year forcing was two orders of magnitude smaller than this (50m/4years). Under these boundary conditions, the various models documented in Griffies et al. (2008) behaved quite differently, with some groups favoring stronger restoring to stabilize the Atlantic overturning.

Here is a summary of some further points to keep in mind regarding salinity forcing.

- At GFDL, we use a real water flux instead of a salt flux. The salinity restoring may be converted to a water flux, or may remain as a salt flux. In the original simulations documented in Griffies et al. (2008), the salinity restoring was converted to water flux. Recent experiment retain the salinity restoring as a salt flux. The preference for salt flux is simply to maintain diagnostic control over the total water budget arising from P-M+R, and to not have that budget confused with added water from restoring.

- To ensure that there is no accumulation of salt in the model arising from the salinity restoring, it is useful to remove the globally integrated salt content from the restoring field at each model time step. When running with real water fluxes, this normalization occurs on the precipitation minus evaporation implied by the salinity restoring.

- As the ocean SST will deviate from that used to balance the dataset's water content, there is no guarantee that the water will balance as the model integrates. Hence, in addition to removing the global mean salt/water associated with the restoring, we remove the global mean evaporation minus precipitation minus river runoff that results from the bulk formulae. Again, this normalization ensures that no water accumulates in the model, and the normalization is applied at each model time step.

5 Closing remarks

As of May 2008, the CORE-IAF.v2 has now released. We anticipate far more efforts at GFDL, and elsewhere, to use this dataset for retrospective simulations in support of reanalysis projects, as well as further model-model comparison projects. We will report on these efforts when further work has been completed.

References

Doney, S. C., S. Yeager, G. Danabasoglu, W. Large, and J. McWilliams, 2003. Modeling global oceanic inter-annual variability (1958-1997): Simulation design and model-data evaluation. NCAR Technical Note: NCAR/TN-452+STR, pp. 48. CGD Division of the National Center for Atmospheric Research.

Doney, S. C., S. Yeager, G. Danabasoglu, W. Large, and J. McWilliams, 2007. Mechanisms governing interannual variability of upper-ocean temperature in a global ocean hindcast simulation. *Journal of Physical Oceanography*, **37**, 1918–1938.

Griffies, S. M., A. Biastoch, C. Böning, F. Bryan, E. Chassignet, M. England, R. Gerdes, H. Haak, R. W. Hallberg, W. Hazeleger, J. Jungclaus, W. G. Large, G. Madec, B. L. Samuels, M. Scheinert, A. S. Gupta, C. A. Severijns, H. L. Simmons, A. M. Treguier, M. Winton, S. Yeager, and J. Yin, 2008. Coordinated Ocean-ice Reference Experiments (COREs). *submitted to Ocean Modelling*.

Large, W., and S. Yeager, 2004. Diurnal to decadal global forcing for ocean and sea-ice models: the data sets and flux climatologies. NCAR Technical Note: NCAR/TN-460+STR. CGD Division of the National Center for Atmospheric Research.

Large, W. G., and S. Yeager, 2008. The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, **submitted**.

Röske, F., 2001. An atlas of surface fluxes based on the ECMWF reanalysis: A climatological dataset to force global ocean general circulation models. Report No. 323. Max-Planck-Institut für Meteorologie.

Zhang, S., R. Greatbatch, and C. Lin, 1993. A re-examination of the polar halocline catastrophe and implications for coupled ocean-atmosphere models. *Journal of Physical Oceanography*, **23**, 287–299.